

CHAPTER 3

DATA ELEMENT DESIGN, DEFINITION, AND NAMING

A. PURPOSE

This chapter provides guidance for designing, defining, and naming data elements that can be used throughout the Department by multiple functional communities. Implementation of this guidance is covered under the processes discussed in Chapters 4, 5, and 6, below.

B. DATA ELEMENT DESIGN

The quality of the data element is the key to the sound foundation for all data structures. Proper emphasis on the creation, naming, and definition of data elements will improve the quality of the entire data structure. Standard data elements should be based upon the data entities and data entity attributes identified in the DoD data model, or recommended for expansion of the DoD data model from a lower level data model, to ensure maximum shareability and interoperability of data throughout the Department of Defense. Several considerations are important to the quality of the data element.

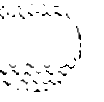
1. Data elements must be designed:

a. To represent the attributes (characteristics) of data entities identified in data models. A model-driven approach to data standards provides a logical basis for and lends integrity to, standard data elements.

b. According to functional requirements and logical, and not physical, characteristics. Physical characteristics include any connotations regarding technology (hardware or software), physical location (databases, records, files, or tables), organization (data steward), or application (systems, applications, or programs).

c. According to the purpose or function of the data element rather than how, where, and when the data element is used or who uses it. It indicates what the data element represents and ensures common understanding.

d. So that it has singularity of purpose. Data elements must not have more than one meaning. A data element should reflect a single concept to promote shareability



and data independence from applications using the data element.

e. With generic element values (domain) that are mutually exclusive and totally exhaustive when the class word “Code” is used.

2. Data elements should not be designed with:

a. Value (domain) that may be confused with another value in the same domain. For example, mixing similar numbers and letters such as 0/0, 1/1, 2/Z, U/V and 5/S.

b. Values (domain) that have embedded meaning or intelligence within part of the code when the class word “Code” is used. For example, do not develop a multiple-character code wherein the value of one or more of the characters in the code have special meaning (i.e., a benefits plan code such as “201,” “202,” “204,” or “205,” where the last digit identifies a particular option within the benefit plan).

c. Overlap or redundancy among the purpose or use of different data elements (e.g., “Birth Date,” “Current Date,” and “Age”).

cm DATA ELEMENT DEFINITION

The definition and naming of a data element is an iterative design process with the data element definition often being modified as the data element is being developed.

1. Data element definitions must:

a. Be based on the definitions of data entity attributes established in the DoD data model or established in an approved data model linked (mapped) to the DoD data model.

b. Have a structure which centers around the generic element of the data it describes. Developing a standard data definition using a structure minimizes “writer’s block” and facilitates the development of consistent and meaningful definitions that can be accepted by all users.

c. Define WHAT the data is, not HOW, WHERE, or WHEN data are used or WHO uses the data.

d. Be more than just a reiteration of the data element name. The definition must add meaning to the name and not merely rephrase the name. The class word is an exception, its meaning does not need to be redefined in each definition.

e. Describe its purpose and usefulness and must not contain physical characteristics. The definition must describe logical, not physical, qualities.

f. Have one and only one interpretation and must not be ambiguous. Terms with differing or varying connotations must have their meanings clearly explained in the definition.

2. Data element definitions must not:

a. Contain conjunctions or phrases indicating multiplicity of purpose of a data element, *ambiguity* of definition, or process orientation.

b. Contain technical jargon that maybe unfamiliar to the reader.

c. Contain acronyms and abbreviations.

d. Restate the characteristics of the data element. For example, do not use statements or phrases such as "... seven characters in length.. ." or "... an alpha-numeric code..." in the definition.

e. Restate a process definition that describes how a data element is calculated, derived, assimilated, or manipulated.

f. Contain information about the valid values or domain of the data elements.

g. Be circular. A situation cannot exist where one definition points to a second definition for further explanation and the second definition points back to the original definition.

D. DATA ELEMENT NAMING

The set of guidelines for naming data elements establishes a naming convention, or classification scheme, that will make it easier to determine if a data requirement is already being met within the Department of Defense or if it is a new requirement that needs to be fully defined and the data collected and distributed as necessary.

1. The names of data elements should:

a. Be based on the names of data entity attributes identified in the DoD data model or an approved data model linked (mapped) to the DoD data model.

b. Be 'clear, accurate, and self-explanatory.

c. Be named according to logical, and not physical considerations. Physical characteristics include any connotations regarding technology (hardware or software), physical location (databases, files, or tables), organization (data steward), or function (systems, application, or programs).

d. Consist of the minimum number of words that categorize the data element. Fewer words may be too general while more words maybe too narrow or restrictive. Modifiers may be used with class words, generic elements, and prime words to fully describe generic elements and data elements. Modifiers are often derived from the data entity attribute names and the entity names identified in the DoD data model or an approved model linked (mapped) to the DoD data model.

e. Include only alphabetic characters (A-Z, a-z), hyphens (-), and spaces().

f. Have each component of the name separated by a space.

g. Have multiple word prime words connected with hyphens. Examples of multiple prime words might be "Purchase-Order," "Medical-Facility," or "Civilian-Government."

2. The following are not permitted in data element names:

a. Words which redefine the data element or contain information that more correctly belongs in the definition.

b. Class words used as modifiers or prime words.

c. Abbreviations or acronyms. (Exceptions to this rule maybe granted by the DoD DAd in the case of universally accepted abbreviations or acronyms. The DDRS will contain a list of approved abbreviations and acronyms.)

d. Names of organizations, computer or information systems, directives, forms, screens, or reports.

e. Titles of blocks, rows, or columns of screens, reports, forms, or listings.

f. Expression of multiple concepts, either implicitly or explicitly.

g. Plurals of words.

h. The possessive forms of a word, i.e., a word which denotes ownership.

L Articles (e.g., a, an, the).

j. Conjunctions (e.g., and, or, but).

k. Verbs.

L Prepositions (e.g., at, by, for, from, in, of, to).